

1 PORTFOLIO THEORY METHOD OF MANAGING OPERATIONAL
2 RISK WITH RESPECT TO NETWORK SERVICE-LEVEL
3 AGREEMENTS

4 CROSS REFERENCE TO RELATED APPLICATION

5 This application claims priority to co-pending U.S.
6 provisional application no. 60/162,383 filed October 28, 1999.

7 Field of the Invention

8 This invention relates to a method of managing risk, and
9 more particularly, to a method of managing operational risk and return
10 with respect to network service-level agreements ("SLA"s).

11 Background of the Invention

12 In order to ensure economical network operations, providers
13 are concerned with the following trade-off: on the one hand, better
14 Quality of Service corresponds to higher price, thereby increasing
15 revenue. On the other hand, if the provider guarantees higher Quality of
16 Service and is not willing to run a higher risk, he can only accept less
17 traffic, thereby decreasing revenue. In order to properly evaluate this
18 trade off, the provider attempts to manage operational risk associated
19 with non-complying network service-level agreements.

20 In the prior art, operators employ simple traffic engineering
21 to meet the QoS as specified in the SLAs. For example, sensitivity
22 analysis is carried out to determine the likelihood of violating SLAs.

1 features and advantages of the present invention will be more fully
2 appreciated by reference to the following detailed description of the
3 presently preferred but nonetheless illustrative embodiments in
4 accordance with the present invention when taken in conjunction with
5 the accompanying drawings.

6 FIG. 1 is a flow diagram of the method of the invention.

7 FIG. 2 is a schematic diagram of a computer device on
8 which the invention operates.

9 FIG. 3 is a diagram of a network on which the invention
10 may be implemented.

11 FIG. 4 is a detailed flow diagram of the method of the
12 invention.

13 FIG. 5 is a schematic view showing portfolio theory applied
14 to network operations.

15 FIG. 6 is a flow chart illustrating a Portfolio Evaluator of
16 the invention.

17 FIG. 7 is a graph of risk vs. return showing the efficient

1 frontier.

2 FIG. 8 is a schematic diagram of an SLA.

3 FIG. 9 is a graph showing examples of extremal points.

4 FIG. 10 is a graph of a polyhedron of constant return.

5 FIG. 11 is a schematic view of an example ring network.

6 FIG. 12 is a graph of the normalized traffic distribution X.

7 FIG. 13 is a zero-profit price curve.

8 FIG. 14 is a graph of the risk and return of portfolios.

9 Detailed Description of the Invention

10 Glossary of Terms and Symbols

11 βC financial penalty per capacity unit.

12 C capacity C (of stem, network, link and so on)

13 \mathbf{D} vector of Quality of Service classes, in case of delay

14 D_i Quality of Service offered by class i in case of delay

15 $e(y)$ return of a portfolio y

- 1 \underline{L} vector of Quality of Service classes, in case of loss ratio
- 2 L_i vector of Quality of Service classes in case of loss ratio
- 3 $p_c C$ constant term reflecting the marginal cost of providing the
- 4 network.
- 5 P denotes a portfolio
- 6 \underline{p} price vector
- 7 p_c unit price for capacity C
- 8 p_i price of contract of type i (expected revenue)
- 9 \underline{q} contracted Quality of Service of the contracts of the portfolio;
- 10 \underline{q}' (expected) actual Quality of Service of a network
- 11 $r(\underline{y})$ risk of a portfolio \underline{y}
- 12 $r_{QoS}(\underline{y})$ Quality of Service risk, i.e., risk expressed in terms of QoS units
- 13 $r_s(\underline{y})$ financial risk, i.e., risk expressed in terms of monetary units
- 14 \mathbf{R} rational numbers
- 15 \mathbf{R}_+ rational numbers that are greater than 0
- 16 \mathbf{R}^n n -dimensional space, where each dimension is of \mathbf{R}
- 17 \underline{y} a portfolio, i.e., $\underline{y} = \langle y_1, \dots, y_i, \dots, y_n \rangle$
- 18 y_i amount of contracts (SLAs) of type i

19 Referring now to FIG. 1, which is a flow diagram of the
20 invention, the invention provides a method 10 and a system 20 that
21 applies the principals set forth in detail in provisional application no.
22 60/162,383, hereby incorporated by reference. The method 10 manages
23 operational risk and return with respect to a portfolio of classes of

1 computer resource or service-level agreements ("SLA"s) by executing
2 the following steps. In a first step 12, the method 10 calculates an
3 efficient frontier 110 that identifies efficient portfolios of SLAs using
4 inputs such as characteristics of the production infrastructure 138, traffic
5 and QoS characteristics and the price of each class of SLAs. In a second
6 step 14, the method 10, optionally, calculating a baseline efficient
7 frontier 110 using inputs such as market pricing and zero-profit pricing.
8 In a third step 16, the method 10 determines the performance of the
9 current portfolio of SLAs using a portfolio evaluator 144 and inputs that
10 characterize the current portfolio. In a fourth step 18, the method 10
11 evaluates performance by comparing the current portfolio and the
12 efficient portfolios with the desired level of risk and return; and, if
13 desired, implements corrective action based on any desired risk and
14 return.

15 Referring now to FIG. 2, which is a schematic diagram of a
16 typical system 20 for practicing the various embodiments of the present
17 invention, the method 10 is encoded on a computer-readable medium
18 and operates on a computer system 20 and/or between the computer
19 system and a server 25 or 54 (shown in FIG. 3) on an intranet or the
20 Internet. Such a computer system 20 typically includes a computer 22, a
21 display device 24, an input device 26 such as a keyboard, a primary
22 storage device 30 and a secondary storage device 32. After loading of
23 software encoded with the method 10 of the invention or after accessing
24 the server 25 or 54 through a browser such as Internet Explore 5.0, as the
25 case may be, the display device 24 displays a graphical user interface

1 ("GUI") 34 for facilitating the display of text and graphics associated
2 with the method to the user. Display devices 24 include printers and
3 computer display screens such as a CRT, LED displays, LCDs, flat
4 screens, screen phones, and projectors. Input devices 26 are numerous
5 and include keyboards and pointing devices such as a mouse 27 having a
6 left mouse button 28 and a right mouse button 29, a trackball, lightpens,
7 thumbwheels, digitizing tablets, microphones using voice recognition
8 software, and touch screens and pads.

9 The GUI 34 provides input fields for data input and control
10 of the method 10, as well as an output window for statistical displays of
11 information, which facilitates management of the network. The method
12 10 accesses a database in primary storage 30, the database including
13 information associated with each SLA, organized in a data structure
14 including the class i of the SLA, the terms 126 of each SLA, such terms
15 including the offered capacity 122, the Quality of Service guarantees 124
16 with respect to delay, loss, and availability, a price 126, a penalty 130, a
17 duration 132, and, optionally, relative compliance guarantee(s) 86a
18 (shown in FIG. 8).

19 The computer 22 includes a CPU 36 as well as other
20 components with which all who are skilled in the art are familiar. For a
21 detailed discussion of these components and their interaction, see U.S.
22 Pat. No. 5,787,254, the content of which is incorporated by reference.
23 The secondary storage 32 supports the method 10, preferably
24 HTTP-compliant, as well as a number of Internet access tools. The
25 CPU 36 fetches computer instructions from primary storage 30 through

1 an interface 40 such as an input/output subsystem connected to a bus 42.
2 The computer 22 can be, but is not limited to, an "IBM APTIVA"
3 computer, a product of International Business Machines Corporation of
4 Armonk, New York, or any computer compatible with the IBM PC
5 computer systems based on the X86 or Pentium(TM) series processor of
6 Intel Corporation or compatible processors, or any other suitable
7 computer. The CPU 36 utilizes an operating system that, depending on
8 the hardware used, may be DOS, "WINDOWS 3.X", "WINDOWS
9 XXXX", "NT", "OS/X", "AIX", "LINUX", or any other suitable
10 operating system. The CPU 36 executes these fetched computer
11 instructions. Executing these instructions enables the CPU 36 to retrieve
12 data or write data to the primary storage 30, display information, such as
13 the statistical displays of the method 10, on one or more display devices
14 24, receive command signals from one or more input devices 26, or
15 transfer data to secondary storage 32 or even other computer systems
16 which collectively form a computer network 25 (shown in FIG. 3).
17 Those skilled in the art understand that primary storage 30 and
18 secondary storage 32 can include any type of computer storage including
19 RAM, ROM, application specific integrated circuits ("ASIC") and
20 storage devices that include magnetic and optical storage media such as
21 a CD-ROM.

22 Where the method 10 operates on a stand-alone computer
23 22, the primary storage 30 stores a number of items including the
24 method 10 and a runtime environment 46. The runtime environment 46
25 typically is an operating system that manages computer resources, such

1 as memory, disk or processor time, required for the method of the
2 invention to run. The runtime environment 46 may also be a message
3 passing system, a microkernel, dynamic loadable linkable module(s), or
4 any other system that manages computer resources.

5 Now referring to FIG. 4, in which a more detailed flow
6 diagram of the method is shown, the method 10 includes the following
7 steps. In a first step 60, the method gathers inputs from the provider
8 including characteristics of the production infrastructure, the QoS
9 characteristics and price of each possible and reasonable class of SLA.
10 In a second step 62, the method 10 calculates an efficient frontier 110
11 (shown in FIG. 7) that identifies efficient portfolios of SLAs.
12 Optionally, the method 10 substitutes the actual pricing of SLAs with
13 baseline pricing such as market pricing or break-even pricing, in order
14 for the operator to obtain insights regarding the effects of price changes
15 on his risk and return, with respect to the market. In a third step 64,
16 which may run concurrently with the first and second steps 60 and 62,
17 the method 10 gathers inputs characterizing the current portfolio of
18 SLAs and the desired risk and return. In a fourth step 66, which may run
19 concurrently with the first and second steps 60 and 62, the method 10
20 computes the risk and return of the current portfolio using a portfolio
21 evaluator 144 (shown in FIG. 6). In a fifth step 70, the method 10
22 calculates the difference between the optimal portfolio identified by the
23 efficient frontier 110 and the current portfolio. In a sixth step 72, the
24 difference is evaluated. If actual risk and return matches the desired

1 levels, then an acceptable portfolio 74 has been attained and the method
2 waits a period of time ΔT (depicted in the figure by box 76), before
3 restarting the method. Otherwise, in a seventh step 80, if actual risk is
4 higher than desired risk or if actual return is lower than desired return,
5 the method 10 takes corrective action. Corrective action can include
6 adjusting marketing strategy 82, changing the degree of multiplexing,
7 84, defining relative compliance guarantees and running packets through
8 a service discipline which allows transmission on the basis of priority (as
9 defined by the guarantees specified in the SLAs), 86, changing prices ,
10 90, trading different classes of SLAs, 92, and/or reducing the costs of the
11 production infrastructure 94. In a seventh step 96, after an adjustment
12 due to the selected corrective action is made to the production
13 infrastructure, the method 10 takes new inputs, and, with the exception
14 of the corrective action of trading SLAs, 92, the method is re-executed,
15 by calculating a new efficient frontier 110 which is compared with actual
16 performance, calculated by the portfolio evaluator 144, given the new
17 parameters.

18 Portfolio Theory and Service-Level Agreements

19 In calculating the efficient frontier 110, the method 10
20 applies the principles of classical Portfolio theory — to be precise, the
21 pre-CAPM (Capital Asset Pricing Model) version of portfolio theory,
22 which was initially developed by H. Markowitz, W. Sharp and others for
23 portfolios of classes of financial assets (shares, bonds, etc.), to provide a

1 framework in which to describe this trade-off between risk and return for
2 portfolios of classes of SLAs. In the classical application of portfolio
3 theory, it is assumed that there are finitely many assets i .

4 Each SLA in the portfolio specifies a peak rate (e.g., bits per
5 second) and a Quality of Service guarantee (e.g., loss rate). Associated
6 with each portfolio is its return (relative profit) and its risk of violating
7 any of the SLAs. This risk will be referred to as non-compliance risk (the
8 risk that any of the Quality of Service guarantees of the sold SLAs is
9 violated). In contrast to return, risk generally cannot be quantified in
10 monetary terms directly. Quantifying risk in monetary terms requires two
11 steps:

- 12 1. Risk is measured in quantities specific to the asset.
- 13 2. The measured risk levels have to be valued in terms of the
14 contract value (e.g., money-back guarantee) specified in the contracts.

15 In order to separate these two steps and apply different
16 valuation methods, risk and return are treated as independent parameters
17 associated with portfolios.

18 Assuming that the set of attainable portfolios is all
19 nonnegative real numbers \mathbf{R}_+ up to the number n of available assets
20 (which is finite), each portfolio may be associated with two quantities:
21 the (expected) return and the risk. A portfolio is called efficient if it
22 maximizes return at a given risk, or equivalently, minimizes risk at a

1 given return. The efficient frontier 110 is the image in the risk-return
 2 space of the set of efficient portfolios.

3 Treating the contracts, e.g., SLAs, of a service provider as a
 4 portfolio makes it possible to develop decision-support tools for
 5 determining the Quality of Service classes to be offered and for
 6 managing the noncompliance risk — an operational risk resulting from
 7 multiplexing (and other behaviors of a production infrastructure in
 8 operation). One such management strategy, corrective action 86, defines
 9 a new contractual parameter called relative compliance guarantees,
 10 which will be discussed in more detail later, along with a discussion of
 11 the trading of risks.

12 Referring now to FIG. 5, a network service provider sells
 13 connectivity over a network 100. The provider offers n classes of
 14 service level agreements $SLA_1 \dots SLA_n$. Each SLA specifies a
 15 connection, contract duration, traffic descriptors (peak rate, average rate,
 16 burst size, etc.) and Quality of Service guarantees (loss rate, delay, jitter,
 17 etc.). An SLA is normalized to a peak rate (or average rate) of 1 bit/s.
 18 Then the provider's portfolio of SLAs is given by a vector $\underline{y} \in \mathbf{R}_+^n$
 19 whose component y_i is the number of contracts of class i .

20 Associated with each portfolio are two quantities: the
 21 (expected) return $e(\underline{y})$ and the risk $r(\underline{y})$. The return (or profit) of a
 22 portfolio \underline{y} equals $e(\underline{y}) = \sum p_i y_i - p_c C$ where p_i , C and p_c denote the unit
 23 price of class i , the capacity of the network and the unit price of network
 24 capacity respectively. Note that $p_c C$ is a constant term reflecting the

1 marginal cost of providing the network. The unit price p_c depends on C
2 as networks 100 exhibit economies of scale in general. In the case of a
3 single link, C is the link capacity.

4 A portfolio \underline{y} entails a risk of noncompliance $r(\underline{y})$ for the
5 provider that depends on the traffic statistics (i.e., the traffic that is
6 actually sent by consumers under their SLAs within the specified traffic
7 descriptor), as well as on the network topology and capacities. There are
8 different risk measures conceivable (as discussed below).

9 In order to help structure his portfolio \underline{y} of SLAs, the
10 provider must consider as inputs such factors as traffic statistics 102,
11 market information 104, and the structure and behavior of the network
12 100. Then, by evaluating risk and return 106, he may determine the
13 efficient frontier 110 (discussed in detail in connection with FIG. 7).

14 The set of feasible portfolios 112 (shown in FIG. 7) and the
15 prices p_i will be determined by the market demand. Once a network
16 service provider has determined the appropriate risk measure, which may
17 be any risk measure, and has derived a way to compute it, he can think
18 about his operations in the terms of portfolio theory. Doing so enables
19 the provider to (1) decide how many and which types of SLAs to offer
20 (described above); (2) evaluate the efficiency of the current portfolio; (3)
21 compute the efficient frontier 110; (4) quantify risk and return 106 of the
22 current portfolio; (5) derive strategies to move towards a more efficient
23 portfolio, and (6) determine base-line portfolios for (cost-based)
24 zero-profit prices.

Evaluate the efficiency of the current portfolio .

In order to obtain the performance characteristics of the existing production infrastructure 138, for comparison with the efficient portfolio 110 (i.e., the fourth step 18 of method 10, shown in FIG. 1), a Portfolio Evaluator 144 is provided. In addition to portfolio details and the production infrastructure (characterized by the vector \underline{i} , which is fixed here and hence not further discussed), the Portfolio Evaluator 144, shown in FIG. 6, takes a Boolean variable “S”, as input to select between risk measure 136a, “ $r_s(\underline{y})$ ”, and risk measure 136b, “ $r_{QoS}(\underline{y})$ ” (i.e., the provider decides whether he wishes to evaluate the risk of a penalty or the risk of violating a Quality of Service requirement) . The Portfolio Evaluator 144 carries out the following steps:

- (1) A Performance Evaluator 146 is invoked to determine the (expected) actual Quality of Service 150, “ q ”. The Performance Evaluator 146 is a formula (if an analytical performance model exists) or a simulator. Further, the actual details of the infrastructure 138 may be used for determining performance.
- (2) The portfolio risk 136, “ $r(\underline{y})$ ”, is computed based on actual Quality of Service 150, q ’, and the contracted Quality of Service 152, \underline{q} , of the contracts of the portfolio using a particular *risk measure* 136a or 136b.
- (3) The return 134, “ $e(\underline{y})$ ”, is computed according to the formula 154,

1 Determining risk measure 136a, $r_s(\underline{y})$, based on specified
2 penalties is just one method to value the risk measure 136b, $r_{QoS}(\underline{y})$, in
3 financial terms, called a *valuation method*. Alternative methods are
4 conceivable including the use of quantified user satisfaction based on,
5 for instance, surveys and experiments. This satisfaction might depend on
6 the market segment (e.g., business and private customers), so that it
7 would be necessary to assign different values to each such group of
8 contracts. A second alternative is given below.

9 In step 80 of method 10, a provider finding out that his risk
10 136b, $r_{QoS}(\underline{y})$, is not zero — he sometimes violates some SLAs — takes
11 corrective action. In corrective action 92, he may re-engineer his
12 infrastructure including increasing the capacity C or accept the risk and
13 pay penalties, if such are specified, or accept unsatisfied customers.
14 Contracts with particularly high Quality of Service guarantees require
15 more resources to guarantee them. However, these high capacity
16 requirements are offset when portfolio mixes such high Quality of
17 Service requirement SLAs with contracts that offer only a low Quality of
18 Service (e.g., a high loss rate) or a low probability of compliance. This
19 leads either to a higher return, lower risk or lower price (or a
20 combination therefore).

21 Referring now to FIG. 4, in corrective action 86, wherein
22 relative compliance guarantees are used, the method 10 of the invention
23 implements a service discipline 86b which allows the degradation the
24 Quality of Service of a communication flow according the Quality of

1 Service specified in the corresponding SLA. The service discipline 86b
2 is carried out by the network (assuming that it is possible to program the
3 network for this purpose) Thus, the provider offers SLAs that guarantee
4 relative compliance 86a. These relative compliance guarantees 86a are
5 specified in terms of a noncompliance risk measure. In practice, a
6 premium is charged for the higher compliance probability. Compliance is
7 hence a product differentiator — a measure which may become as
8 important as network reliability. Note that a portfolio y containing SLAs
9 with relative compliance guarantees 86a can be evaluated with the same
10 approach to evaluate whether these relative compliance guarantees are
11 met. Therefore, the method 10 provides this new contractual parameter,
12 *relative compliance guarantees* 86a. The contractual parameter is
13 calculated in step 70 of the method 10, in which the difference between
14 the actual and the desired risk is equated to the relative compliance
15 guarantee, which is added as a SLA contractual parameter, to define a
16 lower service level. Making the risk explicit enables new valuation
17 methods that, in particular, take advantage of the willingness of
18 consumers to pay a certain amount for a given risk level.

19 Compute the efficient frontier

20 Referring now to FIG. 7, portfolio theory is concerned with
21 the computation and properties of the efficient frontier 110. Once the
22 efficient frontier 110 has been determined, it is a business decision to
23 select a portfolio on the efficient frontier, depending on the tolerable

1 level of risk or the target return.

2 In order for the provider to gain an insight into where his
3 current portfolio stands with respect to an efficient portfolio that
4 maximizes profit for a given risk, steps 60 and 62 of the method 10 apply
5 the principals of Portfolio Theory to calculate the efficient frontier 110.
6 FIG. 7 shows the return-risk space with the attainable portfolios and the
7 set of efficient portfolios, i.e., the efficient frontier 110. The efficient
8 frontier 110 is defined by a closed-form formula, which is only possible
9 in special cases. Assessing the efficiency of the current portfolio P*
10 requires the computation of the efficient frontier 110. The example
11 shown in the figure consists of three segments: two of them result from
12 pairs of adjacent extremal points (shown in FIG. 9 , identified in a closer
13 analysis of the quasi-linearity of the risk function in Portfolio Theory),
14 and the third consists of portfolios of a single Quality of Service class.

15 It is assumed that return 134 is a linear function (as defined
16 above), equal to the summation of the product of each vector describing
17 the portfolio multiplied by nonnegative coefficients of a price vector
18 associated with each vector describing the portfolio, from which
19 marginal cost (a constant) is subtracted.

20 Risk measures 136 can be characterized as convex and
21 quasi-linear risk functions. A function is called convex if all sublevel
22 sets are strictly convex, which yields the following implication, *Lemma*
23 *1*: If the risk measure is a convex risk function, then for every price
24 vector and risk level, there exists a unique portfolio that maximizes

1 return at a given risk level. The function that describes the efficient
 2 portfolios is continuous in both the certain risk level and in the price
 3 vector. The amount of the asset in the unique portfolio is zero whenever
 4 the price vector associated with that asset is also zero.

5 A risk function r is called quasi-linear if it depends only on
 6 the two quantities, the summation of y_i , the vector description of an SLA
 7 in a portfolio and the summation of the product of loss rate L_i for a
 8 particular SLA i and y_i , for some vector $\underline{L} = (L_1, \dots, L_n) \in \mathbf{R}_n^+$ which
 9 characterizes the quality of each SLA (where the lower loss ratio L_i
 10 corresponds to better quality). Note that instead of $\sum y_i$, any linear
 11 function $\sum M_i y_i$ with positive coefficients $M_i > 0$, could have been used
 12 because the transformation $y_i \rightarrow M_i y_i$, $p_i \rightarrow p_i / M_i$ shows that this is
 13 equivalent to the case where all $M_i = 1$. If risk is expressed in terms of a
 14 special function of c , the inverse of the vector of an asset and the loss
 15 ratio, then the condition that the partial derivative of the special function
 16 with respect to c and the partial derivative with respect to the loss ratio
 17 are less than zero ensures that the risk increases with the aggregate of
 18 assets as well as with the quantity.

19 If n , the number of classes of SLAs, is less than 2, a
 20 quasi-linear risk function cannot be convex in the same sense as
 21 described above. The fact that the special function is convex provides
 22 the best proxy of convexity for a quasi-linear risk function.

23 Referring now to FIG. 9, quasi-linearity has the following
 24 consequence: *Lemma 2*: for a quasi-linear risk function, then (i) the

1 efficient frontier 110 is generated by portfolios consisting of one or two
2 classes of SLAs ; (ii) a portfolio consisting of one SLA i is efficient
3 only if (L_i, p_i) , the loss ratio for the SLA and the unit price for that SLA,
4 constitutes an extremal point on the graph of the price $p(L)$ vs. L , loss
5 ratio shown in FIG. 9 , i.e., it lies on the boundary of the curve
6 representing the convex hull in the graph of price vs. loss ratio(therefore,
7 a portfolio of two SLAs, i , and j , is efficient only if (L_i, p_i) and (L_j, p_j) are
8 adjacent extremal points); (iii), supposing that the special function is
9 convex, then there exists a function that assigns to every price vector and
10 risk level greater than or equal to zero, an efficient portfolio of a certain
11 risk consisting of one or two SLAs; and (iv), for a number of SLAs
12 exceeding 2, a function as in “(iii)” cannot be continuous everywhere.

13 Model 1: Loss

14 Assuming that the Quality of Service is described by a
15 single parameter, the loss ratio L , defined as the proportion of lost bits to
16 sent bits in a given time interval of duration T , the relations developed
17 above can be illustrated with a real world example, Model 1, in which
18 the network consists of a single link of capacity C . This is useful due to
19 the fact that single links are important as access lines (e.g., an xDSL line
20 connecting a customer site with a central office) and hot spots, and will
21 be discussed in further detail below. Further, the method 10 assumes that
22 the network employs a proportional scheduling service discipline 86a
23 which ensures that whenever the aggregate condition, defined by the

1 total lost traffic being less than or equal to the summation of the product
2 of the loss ratio L_j and the random variable, X_j , denoting the traffic sent
3 by customers of class j , holds, the lost traffic for each contract does not
4 exceed the specified loss ratio. Then, assuming further that there exists a
5 random variable Y such that $\mathcal{L}X_i \sim (\mathcal{L}y_i)Y$ and $\mathcal{L}_{L_i}X_i \sim (\mathcal{L}_{L_i}y_i)Y$, where \sim
6 denotes equality in distribution, then the risk function is quasi-linear
7 (depending only on $\mathcal{L}y_i$ and $\mathcal{L}_{L_i}y_i$). Therefore, the conclusions (i) and
8 (ii) of *Lemma 2* hold, and one can conclude that the efficient frontier 110
9 is generated by portfolios consisting of at most two Quality of Service
10 classes L_i, L_j corresponding to adjacent extremal points on the price
11 curve.

12 This is consistent with the findings for simple networks of
13 Kai Cieliebbak and Beat Liver, in their provisional application in which
14 it was shown that the efficient frontier 110 is generated by portfolios
15 consisting of at most two Quality of Service classes, L_i, L_j , corresponding
16 to adjacent (i.e., a line segment joining them is contained in the boundary
17 of S) extremal points on the price curve of FIG. 9.

18 In case a network has conceptually a common queue of
19 packets (with respect to the considered Quality of Service parameter), a
20 proportional scheduling policy (with the above-described property) exists
21 and hence Lemma 2 holds. Many broadcast network protocols have this
22 property, so that someone skilled in the art can develop the required
23 proportional scheduling policy. For example, the implementation of this
24 policy for the CSMA/CD (Carrier Sense Multiple Access/Collision

1 Detection) is described as follows. First, each network node has to carry
 2 out admission control. Second, a network node uses the standard
 3 retransmits protocol for dealing with collisions if the lost traffic for class
 4 i exceeds the contracted loss ratio multiplied by the traffic sent by class i
 5 (i.e., $Z_i > \sum_l y_i$). Otherwise, packets are not retransmitted.

6 For non-broadcast networks, routing must be taken into
 7 account. The results for a single link apply only for special cases in
 8 which a network can be treated as a set of independent links. One way
 9 this can occur is if multiplexing among different flows is prevented.
 10 Another possibility is a highly symmetric topology that makes the
 11 network equivalent to independent Links as shown in FIG. 11. A ring
 12 network 160 consisting of four nodes 160a, 160b, 160c, and 160d, four
 13 links 162a, 162b, 162c, and 162d and two flows 164a and 164b between
 14 nodes 160a and 160c, and between 160b and 160d. Flow 164a is equally
 15 distributed over the two possible paths for Flow 164b, and vice versa.
 16 So, for multiplexing purposes, this network 160 is equivalent to a single
 17 link shared by the two flows 164a and 164b. In this figure, a dotted line
 18 represents aggregate flows X^{ij} . In both paths, all links have the same
 19 capacity c_l . The capacity of the ring network 160 depends on $X^{4,8}$: If $X^{4,8}$
 20 varies between 0 and c_l and it is routed clock-wise, the capacity available
 21 to $X^{1,3}$ and $X^{5,3}$ varies between c_l and 0. Consequently, Z depends on the
 22 traffic situation.

23 Model 1 can be applied to real world networks. Networks
 24 fall into two broad categories: broadcast networks (e.g., Ethernet and

1 token ring) and networks using point-to-point connections. Some
 2 broadcast networks have conceptually a common queue of packets, i.e.,
 3 the shared medium may be treated like a single link. For such networks,
 4 the equations for noncompliance risk with loss guarantees and expected
 5 penalty for loss, given below, apply. In fact, there exists a large
 6 number of broadcast networks that can be modeled as a single link.
 7 These include Carrier Sense Multiple Access (CSMA), CSMA/CD
 8 (Collision Detection) – better known as Ethernet, token buses and rings,
 9 wireless networks, and satellite up-links.

10 Quantify risk and return of the current portfolio

11 In the fourth step 66 of method 10, formulas for risk
 12 measures are called for. Two specific formulas for quasi-linear risk
 13 measures may now be provided. First, the following definitions are
 14 made: $y = \sum_i y_i$; $c = C/y$; $L = (\sum_i y_i/y)$, and the random distributions are
 15 written as $Z = (X - C)^+ \sim (yY - C)^+ = y(Y - c)^+$, $\sum_i X_i \sim (\sum_i y_i)Y = LY$.

16 The probability of noncompliance with loss guarantees equals $PNL(c, L)$
 17 $= P[Z > \sum_i X_i] = P[(Y - c)^+ - LY > 0]$. (1)

18 This measure 136 defines the portfolio risk that is the
 19 probability that some SLA of the portfolio is violated. Here the pair of
 20 variables $(y, \sum_i y_i)$ has been replaced with the equivalent pair (c, L) . The
 21 probability of noncompliance can be computed from this formula once

1 the distribution of Y is known (e.g., from historical data).

2 Making the reasonable assumption that the aggregate
3 penalty for noncompliance is proportional to the lost traffic in excess of
4 the SLAs, $Z - \sum_i X_i$, then the expected penalty for loss equals:

$$5 \quad EPL(c, L) = (\beta C) E[Z - \sum_i X_i], \quad (2)$$

6 for some constant $\beta > 0$, so that (βC) denotes the penalty per capacity
7 unit.

8 Model 2: Delay

9 In this section, a second model, Model 2, is described that is
10 complementary to the previous one, based on the following two basic
11 assumptions, namely, (1) a single link and (2) the Quality of Service is
12 described by a single parameter, the delay D . Assume that the link serves
13 customers of guaranteed delays $D_1 < \dots < D_n$. As in the preceding
14 sections, the service discipline is activated which customers of class i
15 have strict priority over customers of class $j > i$ (head-of-line), but
16 service in progress is not interrupted (i.e., non-preemptive).

17 In contrast to the preceding sections, where a general
18 scaling assumption was sufficient, here a specific traffic distribution
19 must be assumed: customers of class i arrive at Poisson rate λ_i , and the
20 arrival processes are independent of each other. Further, service times
21 are identically distributed and they are independent of each other and of

1 the arrival processes (M/G/1 queuing system).

2 Under the assumptions that the network consists of one link
3 of capacity C , the Quality of Service is described by a single parameter
4 (the delay D) and the assumption in the above paragraph, the expected
5 penalty for delay, $EPD(c,D)$, is a quasi-linear risk function that is
6 convex. Therefore, conclusions (i) and (iii) of *Lemma 2* hold: The
7 efficient frontier 110 is generated by portfolios consisting of at most two
8 Quality of Service classes D_i, D_j corresponding to adjacent extremal
9 points on the price curve. Moreover, there exists a function $v^p(p)$ that
10 assigns to every risk ρ and price vector \underline{p} a portfolio of at most two
11 Quality of Service classes, which is continuous except at price vectors
12 where the set of extremal points changes.

13 The expected penalty for delay, EPD is computed over a
14 time interval from the formula: $EPD(c, L) = \beta \sum \{(\lambda_i / \mu_i)(E[W_i] - D_i)\} =$
15 $\beta \{ (\alpha / (c-1)) - (D/c) \}$, where β is a constant > 0 , $c = 1 / \sum (\lambda_i / \mu_i)$, $D = c \sum$
16 $\{(\lambda_i / \mu_i) D_i\}$, and $E[W_i]$ denotes the expected waiting time (i.e., delay) for
17 class i . Assuming that class i traffic arrives at Poisson rate λ_i , and the
18 arrival process are independent of each other; service times,
19 characterized by service rate μ_i of class i , are independently distributed,
20 and they are independent of each other and of the arrival processes —
21 i.e., an M/G/1 queuing system is assumed. Assuming that the service
22 times for customers of all classes are distributed as a random variable Y
23 of mean μ then $\alpha = (1 + \{Var[Y] / \mu^2\}) / 2$, where $Var[Y]$ denotes the
24 variance of random variable Y . Note that noncompliance is defined here

1 in terms of a penalty for exceeding D_i and a premium for remaining
2 under D_i .

3 Determine base-line portfolios for (cost-based) zero-profit prices

4 In step 62 of method 10, determining base-line scenarios, is
5 useful to provide insights in the economics of a network's operation. The
6 method 10 optionally calculates a base-line efficient frontier (or
7 portfolio), assuming that there exists sufficient demand for all
8 considered Quality of Service classes. This means that \mathbf{R}^n_+ defines the
9 set of attainable portfolios. A provider would most likely wish to
10 determine the base-line efficient frontier first. Then, he can investigate
11 which of these portfolios are probably attainable and compare the
12 base-line prices against markets prices (e.g., to determine which Quality
13 of Service classes to offer).

14 For base lining, the prices p_i can be defined as zero-profit
15 prices at the risk level $EPL(c, L)=0$ — so that profit equals costs — by
16 setting prices proportional to the resource consumption of the services.
17 For this purpose, a provider would calculate for a given risk level ρ and
18 Quality of Service class L , the maximal number of contracts $y_{\rho,L}$ he can
19 accept. This yields the profit $e = p y_{\rho,L} - p_C C$, so that the zero-profit price is
20 $p = p_C C / y_{\rho,L}$. The provider is able to offer QoS types profitable if the
21 zero-profit price is equal or lower than market prices. Note that the
22 reverse does not hold, because multiplexing different QoS classes
23 increases often the network utilization and, in turn, reduces the costs.

1 Multiplexing gains (among different QoS types) result in portfolios with
2 $e(y) > 0$. In case that some zero-profit prices are above the market
3 prices, a portfolio y can be considered if the amount $e(y)$ can be used to
4 reduce the prices of contracts that have zero-profit prices above market
5 prices. If a provider calculates the efficient frontier, he would usually
6 eliminate the portfolios from the frontier where he would expect $e(y) <$
7 0 . The reason is that in case $e(y) < 0$, the network exhibits negative
8 multiplexing gains (i.e., the assuming usage pattern cannot be allocated
9 efficiently), the network is not well suited for offering such
10 combinations of QoS classes and, hence, such combinations should not
11 be offered. A prospective provider might calculate the zero-profit prices
12 (i.e., the prices that cover costs) and the resulting base-line efficient
13 frontier. He could then compare these zero-profit prices of SLAs
14 belonging to efficient portfolios with the market prices: if all zero-profit
15 prices associated with each portfolio are, for instance, above the market
16 prices, the provider is not competitive. For a particular portfolio
17 (assuming no other financial subsidies), the losses of due contracts with
18 zero-profit prices that are higher than market prices have to be
19 compensated by profits due to contracts with lower zero-profit prices
20 than market prices.

21 Derive strategies to move towards a more efficient portfolio

22 Referring again to FIGS. 4 and 7, in order to achieve a more
23 efficient portfolio (depicted by the arrow pointing from the current

1 portfolio P^* to the efficient frontier), several options 80 for corrective
 2 action are possible. In corrective action 84, a service provider might
 3 reduce costs or increase risk. For this purpose, the degree of
 4 multiplexing could be increased or the network capacity C decreased.
 5 Note that it is sometimes possible to increase the multiplexing without
 6 modifying the risk. Such a method is described by Kurz, Thiran, and
 7 LeBoudec, in an article entitled *Regulation of a connection admission*
 8 *control algorithm* in the Proceedings of INFOCOM'99. In corrective
 9 action 82, a provider might adopt a marketing strategy to move towards a
 10 more efficient portfolio. For instance, the price of the low-quality
 11 service could be reduced to increase the number of contracts in this
 12 class. In corrective action 92, providers might trade risks (analog to load
 13 securitization and syndication): a provider can buy and sell contracts to
 14 optimize his portfolio assuming that there exists a market for trading
 15 contracts. For trading risks, the operator determines the number of
 16 to-be-traded contracts of class i , $\Delta_i = y_i - y_i^*$, where y_i^* and y_i denote the
 17 number of contracts of class i in case of the current portfolio and a
 18 desirable (i.e., efficient) portfolio, respectively. If $\Delta_i > 0$, it's necessary
 19 buy Δ_i contracts of class i , and if $\Delta_i < 0$ the provider sells this number
 20 of contracts of class i . Note that trading is a corrective action that leads
 21 to an efficient portfolio (assuming that the necessary trades can be
 22 executed, i.e., that there is adequate supply of SLAs having the
 23 appropriate characteristics and a means of purchasing these SLAs).

24 An advantage of the invention is that it automatically and

